Current Biology, Vol. 12, R846–R848, December 23, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(02)01350-7

DNA Repair: Bioinformatics Helps Reverse Methylation Damage

Dispatch

Josef Jiricny

Recent work has uncovered a novel DNA repair enzyme: the AlkB protein of *Escherichia coli*, which oxidises the methyl groups of 1-methyladenine and 3-methylcytosine to hydroxymethyl moieties; the oxidised groups are subsequently released as formaldehyde, regenerating the unmodified bases.

Exposure of *Escherichia coli* to methylating agents evokes a so-called adaptive response [1], involving the induction of the *ada*, *alkA*, *alkB* and *aidB* genes. Although the expression of any one of these genes protects the bacterial cells — at least partially — against the cytotoxic and mutagenic effects of methylating agents, the biochemical mechanisms underlying this protection have until recently been known only for the Ada and AlkA proteins (see below). AidB's function is still not known; it has been suggested that it prevents methylation by detoxifying nitrosoguanidines [2], but this hypothesis remains to be substantiated. Similarly, it has been known for nearly two decades [3] that AlkB protects against the cytotoxic effects of 'S_N2' methylating agents, such as methyl methane sulphonate (MMS), and it has been shown to efficiently detoxify methylated single-stranded DNA [4], but until now all attempts at elucidating AlkB's mechanism of action failed.

This situation has now been redressed by studies [5,6] which have confirmed the insightful prediction by Aravind and Koonin [7] that AlkB is an α-ketoglutarate- and Fe(II)-dependent dioxygenase. With the help of its cofactors, the AlkB enzyme converts 1-methyladenine and 3-methylcytosine to 1-hydroxymethyladenine and 3-hydroxymethylcytosine, respectively. The subsequent spontaneous release of formaldehyde regenerates unmodified adenine and cytosine. AlkB thus joins the other damage-reversal enzymes, DNA-photolyase [8] and O^6-methylguanine DNA-methyl transferase, in the gallery of proteins that repair DNA damage by direct reversal of the modification, without resorting to cleavage of glycosidic bonds or the sugar-phosphate backbone of DNA.

All organisms studied to date possess sophisticated defense systems that protect them against the deleterious effects of methylating agents. More than one repair pathway is required, because these agents come in many guises and modify DNA bases at most nitrogen and oxygen atoms (Figure 1). The methylating agents can be divided into two classes, S_N1-type and S_N2-type, according to the chemical mechanism of methylation. S_N1-type methylating agents, such as methylnitrosoguanidines and methylnitrosourea — which arise during the decay of organic matter under acidic conditions — modify primarily the N^7 position of guanine, the N^3 position of adenine and the O^6 position of guanine, but form also small amounts of pyrimidines methylated on the O^2 and O^4 positions. S_N2 agents such as dimethylsulphate (DMS), MMS and the naturally occurring methyl chloride, a product of chloride detoxification, generate in addition N^1-methyladenine and N^3-methylcytosine in single-stranded DNA (see [1] for a recent review). Some of these modifications destabilize the bases to hydrolysis, leading to their loss or ring-opening; others alter the bases' hydrogen-bonding properties (Figure 1). Because the biological consequences of these modifications range from blocked replication or transcription to the generation of mutations or DNA breaks, it is imperative that they be repaired in a cell with high efficiency and specificity — as indeed they are.

The base-excision repair pathway, which plays a key role in the removal of most damaged DNA bases, is involved also in the detoxification of methylated DNA. In this repair process, a damage-specific DNA-glycosylase removes the aberrant bases by cleaving the glycosidic bond between the damaged/modified heterocycle and the deoxyribose residue of the DNA backbone [9]. This reaction gives rise to an abasic site in the DNA, which is further processed by an apurinic-endonuclease. The gap that arises through the removal of the baseless sugar–phosphate is filled in by a DNA polymerase, and the remaining nick is sealed by DNA ligase [10].

E. coli has two DNA glycosylases that deal with methylated bases. The product of the constitutively expressed *tag* gene removes N^3-methyladenine from DNA [11], while the inducible *alk*A gene encodes a DNA glycosylase which processes not only the most abundant methylated purines, N^7-methylguanine and N^3-methyladenine, but also N^3-methylguanine, O^2-methylcytosine and O^2-methylthymine [12]. The focus on N^3-methyladenine is necessary, as this modified base — generated even at low concentrations of methylating agents — blocks DNA replication and transcription, unlike N^7-methylguanine, which is formed at higher amounts but which is apparently innocuous.

The major products of methylation of the exocyclic oxygen residues, O^6-methylguanine and O^4-methylthymine, are not substrates for base-excision repair. Instead, they are processed by O^6-methylguanine DNA-methyl transferases, members of the damage-reversal family of repair proteins. The *E. coli* genome includes two genes encoding proteins of this type — one inducible (*ada*) and one constitutively expressed (*ogt*). Under normal conditions, both genes are expressed at very low levels. But when the Ada protein detects a methylated phosphate moiety — a methyl phophotriester — it transfers the methyl group onto one of its internal cysteine residues. This brings about a conformational change in the polypeptide,

Institute of Medical Radiobiology of the University of Zürich and the Paul Scherrer Institute, August Forel-Strasse 7, CH-8008 Zürich, Switzerland. E-mail: jiricny@imr.unizh.ch

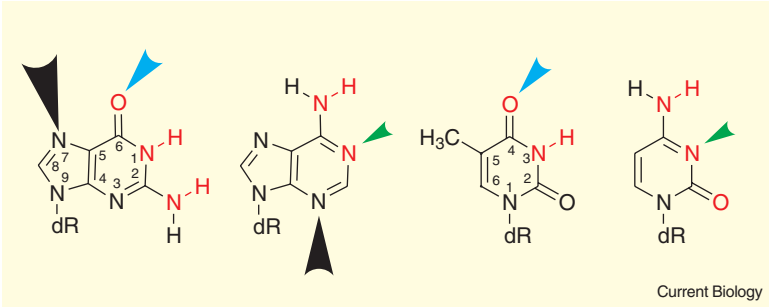

Figure 1. Main sites of modification of DNA bases by methylating agents.

The arrow sizes reflect the relative yields of the different methylated products. Atoms or groups involved in Watson–Crick hydrogen bonding are shown in red. Bases methylated at positions indicated by the black arrows are processed primarily by AlkA; those methylated at sites marked with blue arrows are addressed by O^6-methylguanine DNA methyltransferase, and those methylated at positions marked with the green arrows are processed by AlkB. (See text for details.)

which converts the protein into a transcription factor that activates the expression of the *ada*, *alkA*, *alkB* and *aidB* regulon.

The newly synthesized AlkA protein deals with the principal cytotoxic alterations (see above), while the Ada polypeptide removes methyl groups from the O^6 position of guanine and the O^4 position of thymine. Notably, Ada fails to turn over; it becomes inactivated as a result of the methyl group transfer from the exocyclic oxygen of the base onto its residue cysteine 321 [1] — that is why it has often been termed a suicide enzyme.

The new findings [5,6] show that AlkB, like Ada, is also a damage-reversal protein. But in contrast to Ada, which simply transfers the methyl group from an oxygen atom of the base onto a sulfur atom of one of its cysteines, the AlkB protein requires iron (II), oxygen and α-ketoglutarate as cofactors, and the process of converting a methylated base to an unmodified one involves the oxidation of the methyl group to formaldehyde, the conversion of the ketoglutarate to succinate and CO_2, and the oxidation of Fe(II) to Fe(III) (Figure 2). Why does the organism have to go to such lengths to remove such small modifications and why chose such an unusual strategy? Why not just use base-excision repair, for example?

The answer to these questions may lie in the following considerations. Base-excision repair is restricted to double-stranded DNA substrates. Glycosylase-mediated generation of an abasic site in single-stranded DNA can occur, but this results in irreparable strand cleavage, as no rescue of this lesion is possible in the absence of a complementary template. Because the major sites modified by methylating agents — the N^7 position of guanine and N^3 position of adenine — are not involved in Watson–Crick base-pairing interactions, bases methylated at these positions are formed in large amounts in double-stranded DNA, and can be efficiently repaired in this substrate by the base-excision repair machinery. In contrast, the O^6 position of guanine, the O^4 and N^1 positions of adenine and the N^3 position of cytosine are involved in hydrogen bonding (Figure 1) and are therefore accessible to the methylating agents predominantly in single-stranded DNA.

The exocyclic oxo-groups of guanine and thymine might be modified to a small extent even in double-stranded DNA, because of their ready accessibility from the major groove and their ability to undergo keto-enol tautomerism during DNA 'breathing', which allows methylation of the enol form. But this is not the case for the N^1 position of adenine and N^3 position of cytosine, both of which lie buried deep in the interior of the double helix and which are therefore accessible to reagents only in single-stranded DNA regions, such as replication forks or transcription bubbles.

The dangers associated with attempting to repair these substrates in single-stranded DNA by base-excision repair probably drove the organisms to evolve repair pathways that did not involve scission of the phosphodiester backbone of DNA. In the case of *O*-methylations, damage reversal just required a methyl group acceptor on the part of the repair protein that would be more nucleophilic than oxygen. The sulfhydryl group of cysteine was ideally suited for this task; though the price of this rescue was inactivation of the protein, this could be tolerated, because the *O*-methyl modifications do not block essential processes of DNA metabolism.

But methylation of the N^3-position of adenine and the N^1-position of cytosine arrests replication and transcription, so a different strategy for eliminating these *N*-methylations had to be found, and nature found it. The strategy that evolved is to oxidise the methyl substituent; the resulting hydroxymethyl moiety is a much better leaving group than methyl and can be spontaneously lost from the base. Moreover, this reaction can be catalysed by a real enzyme — AlkB turns over on its substrate [6] — which is active on both single-stranded and double-stranded DNA [5], so the cell can be always protected.

From where did the *alkB* gene of *E. coli* originate? This is not clear, but AlkB is not the only α-ketoglutarate- and Fe(II)-dependent dioxygenase. Rather, it is a member of a very large family of proteins that catalyze a variety of reactions typically involving the oxidation of organic substrates [7]. Database searches identified more than eighty representatives of this family in organisms for which sequence information is available, twenty five of which are AlkB homologs [7]. Three of these may turn out to be AlkB orthologs — one has already been described, even though its biological activity was extremely low [13] — but the function of the others is unclear. One prediction is that they will turn out to be involved in RNA detoxification, possibly in the reactivation of viral RNA that has been modified by host-encoded RNA methylases [7]. But these speculations need to be substantiated experimentally.

In the closing sentence of their paper, Sedgwick and colleagues [5] speculate on the possible biological roles of some of these AlkB homologs. One suggestion

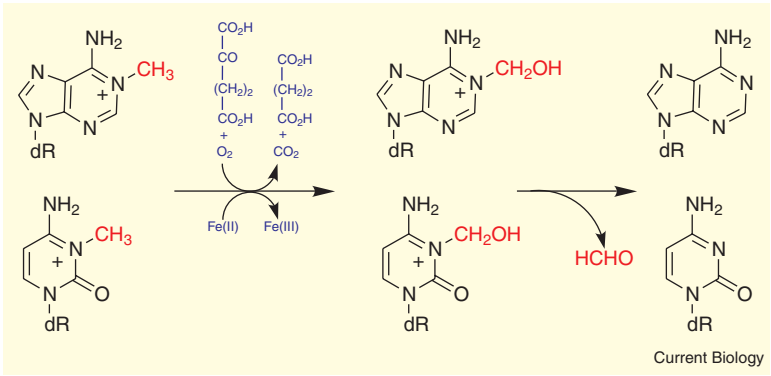

Figure 2. AlkB-mediated demethylation of N^3-methyladenine and N^1-methylcytosine.

The methyl groups are converted to hydroxymethyl moieties, which are spontaneously lost in the form of formaldehyde to regenerate the unmodified bases (the aberrant methyl groups and their oxidation products are shown in red). Note that the cofactors (blue) are consumed during the reaction in stoichiometric amounts.

is that they might participate in the demethylation of 5-methylcytosine, which has been reported to take place on a massive scale in the paternal nucleus of a zygote [14], and – more selectively or sporadically – during differentiation [15] and transformation [16]. The demethylase involved — the current Holy Grail of epigenetics — has so far eluded detection, despite numerous attempts. It has long been argued that the C–C bond that would have to be broken during conversion of 5-methylcytosine to cytosine is too stable to be cleaved. But contrary to these expectations, the methylated DNA-binding protein MBD2 was reported to catalyse just such a cleavage reaction [17,18], and, what is more, to do so by a mechanism similar to that described for AlkB. In contrast to AlkB, however, MBD2 was reported to require no cofactors for the conversion of the methyl group to a hydroxymethyl moiety and then to methanol.

Given the lengths to which AlkB and its homologs have to go to fulfil their respective biological roles, the lack of cofactor requirement in the MBD2-catalysed demethylation is highly surprising. But then, the reported activity of MBD2 was barely detectable, which implies either that the activity is not physiological, or that the right reaction conditions and/or cofactors remain to be identified. The search for the elusive 5-methylcytosine is therefore still open. It will be interesting to see whether it finally does turn out to be an oxygenase, AlkB-related or otherwise.

The functional characterisation of the AlkB proteins of *E. coli* has provided us with a new insight into DNA — and perhaps also RNA — metabolism. The biochemical studies of the Sedgwick and Seeberg groups were impeccable, but my hat goes off to the bioinformatitions, Aravind and Koonin [7], who pointed the biochemists to the correct cofactors. Without them, the DNA repair community might still be going through cupboards full of chemicals in search of the magic ingredients that make AlkB tick.

References

1. Sedgwick, B. and Lindahl, T. (2002). Recent progress on the Ada response for inducible repair of DNA alkylation damage. Oncogene, in press.
2. Landini, P., Hajec, L.I. and Volkert, M.R. (1994). Structure and transcriptional regulation of the *Escherichia coli* adaptive response gene aidB. J. Bacteriol. *176*, 6583–6589.
3. Kataoka, H., Yamamoto, Y. and Sekiguchi, M. (1983). A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate. J. Bacteriol. *153*, 1301–1307.
4. Dinglay, S., Trewick, S.C., Lindahl, T. and Sedgwick, B. (2000). Defective processing of methylated single-stranded DNA by E. coli AlkB mutants. Genes Dev. *14*, 2097–2105.
5. Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T. and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature *419*, 174–178.
6. Falnes, P.O., Johansen, R.F. and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. Nature *419*, 178–182.
7. Aravind, L. and Koonin, E.V. (2001). The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol. *2*, research0007.1–0007.8.
8. Sancar, A. (1996). DNA excision repair. Annu. Rev. Biochem. *65*, 43–81.
9. Scharer, O.D. and Jiricny, J. (2001). Recent progress in the biology, chemistry and structural biology of DNA glycosylases. Bioessays *23*, 270–281.
10. Seeberg, E., Eide, L. and Bjoras, M. (1995). The base excision repair pathway. Trends Biochem. Sci. *20*, 391–397.
11. Bjelland, S. and Seeberg, E. (1987). Purification and characterization of 3-methyladenine DNA glycosylase I from *Escherichia coli*. Nucleic Acids Res. *15*, 2787–2801.
12. Tudek, B., Van Zeeland, A.A., Kusmierek, J.T. and Laval, J. (1998). Activity of Escherichia coli DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. Mutat. Res. *407*, 169–176.
13. Wei, Y.F., Carter, K.C., Wang, R.P. and Shell, B.K. (1996). Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. Nucleic Acids Res. *24*, 931–937.
14. Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.
15. Reik, W., Dean, W. and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089–1093.
16. Jones, P.A. and Laird, P.W. (1999). Cancer epigenetics comes of age. Nat. Genet. *21*, 163–167.
17. Bhattacharya, S.K., Ramchandani, S., Cervoni, N. and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. Nature *397*, 579–583.
18. Detich, N., Theberge, J. and Szyf, M. (2002). Promoter-specific Activation and Demethylation by MBD2/Demethylase. J. Biol. Chem. *277*, 35791–35794.